A Note on a Practical Stochastic Dominance Based Solution to Public Policy Choice When Confronted with a Set of Mutually Exclusive Non-combinable Policy Prospects

Gordon Anderson*                  Teng Wah Leo†
University of Toronto              St. Francis Xavier University

13 April 2014

*Department of Economics, University of Toronto. Email Address: anderson@chass.utoronto.ca
†Department of Economics, St. Francis Xavier University. Email Address: tleo@stfx.ca
Abstract

The problem considered here is that of a policymaker’s choice between policies when facing a set of distinct, non-combinable options. When policies are not combinable, for example when public investments are “lumpy”, and a convex combination of the two options is not feasible, the classic comparative static solution is not available. The approach proposed here is an adaptation of the solution to the problem of choosing the best statistical test amongst a collection of tests, based upon their power function properties. Here, policies are evaluated using a statistic that measures the proximity of the distribution of its respective outcomes to a “stochastically dominant” ideal generated from a synthetic combination of all policies under a policymaker’s given imperative. The paper concludes with an illustrative example.
1 Introduction

A cornerstone in the advance of expected utility and prospect based choice theory when considering an aggregate or expected utility objective (Kolm (1966), Atkinson (1987), Foster and Shorrocks (1988), Rothschild and Stiglitz (1970), and Kahneman and Tversky (1979)), has been the use of stochastic dominance criteria (modified in the case of prospect theory) to establish the unambiguous superiority of one outcome distribution over another. Predicated upon the nature of preferences, it provides a set of conditions which the preferred outcome distribution associated with a particular state of the world should satisfy relative to its competitor state. The technique has a wide range of applications, yet in spite of a well-developed theory for public policy application (Moyes and Shorrocks 1994), it seems to be seldomly used in practice by policymakers (or portfolio managers). The primary reasons being firstly at low orders of dominance the method only provides an incomplete ordering (comparisons are not always conclusive), and secondly the method does not yield a measure of “by how much one policy (or portfolio) is better than another” for policymakers (or investors) to hang their hats on in the choice process.

These issues associated with the practicality of stochastic dominance techniques have long been recognized in the finance literature, where the conventional second order stochastic dominance criterion appropriate for risk averse actors (Rothschild and Stiglitz 1970) was acknowledged by the same authors (Rothschild and Stiglitz 1971) to have no obvious comparative statics properties\footnote{They demonstrated that an increase in risk characterized by 2\textsuperscript{nd} order dominance does not necessarily induce all agents to reduce holdings of the risky asset.}. In this literature, a response to this concern was the introduction of an alternative form of dominance, namely “central dominance”, characterizing “greater central riskiness” (Gollier 1996). Central Dominance, which is neither stronger nor weaker than second order stochastic dominance, characterizes the necessary and sufficient conditions under which a change in risk changes the optimal value of an agent’s decision variable in a predictable fashion for all risk-averse agents\footnote{Chuang et al. (2013) have developed tests for Central dominance.}. An important feature of this analysis is that the decision variable is continuously related to the risk measure\footnote{The risk free–risky asset mix parameter in the case of the portfolio problem or the tax parameter(s) in a public choice problem.}, so that incremental changes in the decision variable can be contemplated as a consequence of incremental changes in risk. However, the notion of Central Dominance has not yet found expression in the wellbeing policy literature (for an exception see...
Chuang et al. (2013)), probably because in many situations policy alternatives are generally not continuously connected in the manner that a convex combination of a risky and risk free asset can be contemplated in the portfolio problem\(^4\). Rather, policy alternatives are a collection of distinct, non-combinable policies or prospects, and the choice problem is that of picking one of them. In these circumstances, where a variety of alternative policy outcomes is being contemplated (usually in terms of the income distributions they each imply), a collection of pairwise dominance comparisons will have to be made without recourse to the comparative static feature that the notion of central dominance provides.

While much can be learned about the relative status of alternative policy outcomes by considering them under different orders of dominance comparisons, the partial ordering nature of the technique frequently renders the comparisons inconclusive. The usual practice in the empirical wellbeing literature is to compare alternative outcome distributions (usually income size distributions) at successive orders of dominance, until dominance at a given order is established. Unfortunately in terms of a collection of pairwise comparisons, this can be a complicated and lengthy process which is frequently impractical (hence the lack of its use). In fact successive orders of dominance comparison attach increasing importance (weight) to lower values of the income variable in question, so that increasing orders of dominance may be construed as reflecting “successively increasing degrees of concern for the poor” policy imperatives that confront a policymaker.

Here indices are proposed for measuring the extent to which one policy is “better” than another within the context of a specific dominance class, the choice of which reflects the particular imperative confronting the policymaker. Conceptually the index is based upon the approach taken in the statistics literature\(^5\) to choosing from a finite collection of alternative tests, on the basis of the eyeballed proximity of each test’s power function to the envelope of the set of available power functions. The envelope reflects the maximal power that could be obtained if the best bits of each test could be notionally combined. So here alternative policy options are considered in the context of a dominance class determined by the policymaker’s imperative. The stochastically dominant envelope of policy consequences at the given order of dominance is constructed, so that a measure of the proximity to this envelope for each of the policy options can be calculated. The policy option most proximate to this dominating envelope, i.e. the one with the smallest

\(^4\)Central Dominance could for example be employed in examining a revenue neutral redistributive tax policy, which is some convex combination of lump sum and progressive tax.

\(^5\)See for example Ramsey (1971), Juhl and Xiao (2003), and Omelka (2005).
proximity measure is to be preferred.

In the following, Section 2 outlines the relationship between stochastic dominance criteria, wellbeing classes and the notion that a policymaker may want to make a policy choice in the context of an imperative associated with a particular wellbeing class. In Section 3 the indices appropriate for making such choices are developed. Section 4 exemplifies the technique using a sample of weekly pre-tax incomes drawn from the Canadian Labour Force Survey for January 2012, as the basis for three non-combinable alternative revenue neutral policies from which a choice has to be made.

2 Stochastic Dominance, Wellbeing Classes and the Policymaker’s Imperative

The notion of stochastic dominance was developed as a criteria for choosing between two potential distributions of a random variable \( x \) (usually income, consumption or portfolio returns) in order to find the distribution which maximizes \( \mathbb{E}(U(x)) \) based upon the properties of the function \( U(x) \), where \( U(x) \) represents a felicity function of agents in a society under the income size distribution of \( x \) (Levy (1998) provides a summary). The technique yields a decision as to which is the preferred distribution at some specification of \( U(x) \), where successively restrictive specifications of \( U(x) \) require successively higher orders of dominance comparison. Since dominance at lower orders of comparison implies dominance at higher orders of comparison, the practice is to start comparison at the first order, and making comparison at successively higher orders until an unambiguous decision is reached.

Working with \( U(x) \), let \( x \) be continuously defined over the domain \([a,b]\), and two alternative states defined by density functions \( f(x) \) and \( g(x) \) describing the distribution of \( x \) across agents in those states. The family of Stochastic Dominance techniques address the issue: “which state is preferred if the objective is the largest \( \mathbb{E}(U(x)) \)” . Formally, when the derivatives of \( U(x) \) are such that \((-1)^{i+1} \frac{d^i U(x)}{dx^i} > 0\), for \( i = 1, \ldots, J \), a sufficient condition for:

\[
\mathbb{E}_f[U(X)] - \mathbb{E}_g[U(X)] = \int_a^b U(x)(dF - dG) \geq 0
\]

\(6\)Sometimes poorness or poverty indices \( P(x) \) are studied (Atkinson 1987) in which case \( U(x) = -P(x) \).
is given by the condition for the dominance of distribution $G$ by $F$ at order $j = 1, 2, \ldots, J$, which is:

$$F_j(x) \leq G_j(x) \quad \forall \ x \in [a, b] \quad \text{and} \quad F_j(x) < G_j(x) \quad \text{for some} \ x \in [a, b] \quad (2)$$

where: $F_i(x) = \int_{a}^{x} F_{i-1}(z)dz$ and $F_0(x) = f(x)$

Essentially the condition requires that the functions $F_j(x)$ and $G_j(x)$ not cross, so that the dominating distribution is “unambiguously” below the other. It will be useful for the subsequent discussion to note that $F_i(x)$ (or equivalently $G_i(x)$) may be rewritten in incomplete moment form as:

$$F_i(x) = \frac{1}{(i-1)!} \int_{0}^{x} (x - y)^{i-1}dF(y) \quad (3)$$

An important notion regarding dominance relations in what follows is that dominance at order $h$ implies dominance at all orders $h' > h$, and a useful lemma in Davidson and Duclos (2000) is that if $F$ first order dominates $G$ over some region $(-\infty, a)$ then $F$ will dominate $G$ over the whole range of $x$ at some higher order. Thus the practice has been to seek the order at which dominance of one distribution over the other is achieved, for such a comparison is unambiguous at that order of dominance. There is also the implication that at a sufficiently high order of dominance the ordering will be complete rather than partial.

From (2) it may be seen that the dominating distribution is the preferred distribution, reflecting as it does the desire for greater expected $U(x)$. However, as can be seen from (3) increasing orders of dominance attach increasing weight to lower values of $x$ in the population distribution. Thus successively higher orders of dominance can be interpreted as reflecting higher orders of concern for the “poor” end of the distribution. Following Foster and Shorrocks (1988) this permits the interpretation of various forms of dominance as follows$^{7}$:

- $U_{i=1}$, which only requires $\frac{dU(x)}{dx} > 0$, and yields a $1^{st}$ order dominance rule, is referred to as *Utilitarian* societal preference, and is really an expression of preference for more

$^{7}$The comparison procedures have been empirically implemented in several ways, see for example Anderson (1996, 2004), Barrett and Donald (2003), Davidson and Duclos (2000), Linton et al. (2005), Knight and Satchell (2008), and McFadden (1989).
of $x$ without reference to the spread of $x$. In the context of the dominance relation, the weight attached to each value of $x$ in the population distribution is the same. However in the cumulative distribution, heuristically the first increment of $f(x)$ is counted at every value of $x$, the second increment of $f(x)$ is counted at every value of $x$ except the first, the third increment of $f(x)$ is counted at every value of $x$ except the first and second . . . . In terms of the policymaker’s imperative, she would be indifferent to revenue neutral transfers between agents.

- $U_{i=2}$, which requires $\frac{dU(x)}{dx} > 0$, $\frac{d^2U(x)}{dx^2} < 0$, and yields a 2nd order dominance rule, is referred to as Daltonian societal preference, and is an expression of preference for more $x$ with weak preference for reduced spread. On the margin, for two distributions with equal means but different variances (whose cumulative distribution will cross, thus contradicting the 1st order dominance criteria), the one with the smallest variance will be preferred. In the context of the dominance relation formula, the weight attached to each value of $x$ in the population distribution decreases as $x$ increases (heuristically the first increment of $f(x)$ is counted twice at every value of $x$, the second increment of $f(x)$ is counted twice at every value of $x$ except the first, the third increment of $f(x)$ is counted twice at every value of $x$ except the first and second, . . . , and in terms of $x$, equation (3) reveals the increments are units of $x$.

In terms of the policymaker’s imperative she would have a preference for revenue neutral transfers from rich agents to poor agents.

- $U_{i=3}$, which requires $\frac{dU(x)}{dx} > 0$, $\frac{d^2U(x)}{dx^2} < 0$ and $\frac{d^3U(x)}{dx^3} > 0$, yields a 3rd order dominance rule, and is an expression of preference for more, with a weak preference for reduced spread especially at the low end of the distribution. In the context of the dominance relation, the weight attached to each value of $x$ in the population distribution decreases at an even faster rate as $x$ increases (heuristically the first increment of $f(x)$ is counted thrice at every value of $x$, the second increment of $f(x)$ is counted thrice at every value of $x$ except the first, the third increment of $f(x)$ is counted thrice at every value of $x$ except the first and second, . . . , and in terms of $x$, equation (3) reveals the increments are squares of units of $x$).

In terms of the policymaker’s imperative she would have a preference for revenue neutral transfers from rich agents to poor agents, and the preference would be stronger the poorer the agent.

...
• $U_\infty$ or infinite order dominance is referred to as Rawlsian societal preference, since it attaches infinite weight to the poorest individual, and can be examined in the context of the relative incomes of the poorest individuals in two equally populated societies. Essentially the outcome distribution which yields the best outcome for the poorest individual is the one that is chosen.

With this in mind, the policymaker is lead to contemplate a particular order of dominance (choice of $i$) in order to reflect the imperative she confronts, in terms of the degree of concern for the poorer agents in a society. Thus if the policymaker was indifferent as to where in the distribution of incomes revenue neutral transfers were made, she would consider 1st Order Dominance comparisons. If on the other hand the policymaker deems it politic to give added weight to the concerns of the poor, policy comparisons should be conducted in terms of higher orders of dominance (values of $i$ greater than 1) of the distributions of policy outcomes. In this context, it may be that no policy dominates at the chosen level of concern. However the policymaker could choose the policy which gets closest to the envelope of alternative policy outcomes, at the appropriate level of integration ($i$) which reflects the imperative she confronting, if indices of proximity to the envelope were available.

### 3 The Comparison Indices

Suppose we are to contemplate a collection of wellbeing distributions $G(x)$, $H(x)$, $J(x)$, $\ldots$, $K(x)$, which are the consequence of alternative policy measures, where for convenience $x \in [0, \infty)$. In the context of wellbeing comparisons, we are lead to consider a collection of pairwise comparisons within the family of dominance criteria where $j$’th order dominance is of the form given in (2) above. Anderson (2004) interpreted dominance between $F$ and $G$ at a particular order as a measure of the degree of separation between the distributions at that order, and the area between the two curves provides a very natural measure of the magnitude of the separation. Furthermore, when $G$ dominates $F$ at the $j$’th order, and given $\mu^j$ is a location measure of $x$ such as the mean, median or modal value of $x$, note then that:

$$\mathcal{PB}_j = \frac{-1}{\mu^j} \int_{0}^{\infty} [G_j(z) - F_j(z)] \, dz$$

(4)
provides a standardized measure of such a separation or wellbeing excess of \( G \) over \( F \), where the metric of the unstandardized measure is related to the units of \( \mu^j \), making this a unit free measure. However such an index only works if \( G \) dominates \( F \) at this order. What if there is no dominant policy at a given order?

Suppose the policymaker’s imperative is utilitarian, it may well be that there is no 1st order dominant policy in the collection. Consider the lower frontier or envelope of all distributions \( G(x), H(x), J(x), \ldots, K(x) \) in the collection given by \( LE(x) = \min\{G(x), H(x), J(x), \ldots, K(x)\} \). Thus in so doing, effectively the “best policy” of each point \( x \) has been selected to produce the best possible synthetic policy over the whole range of \( x \), if all policies could be combined. Obviously \( LE(x) \) would dominate all distributions \( G(x), H(x), J(x), \ldots, K(x) \) at the first order and would thus, if it existed, be the preferred distribution (Note that if one of the distributions in the collection 1st Order Dominated all of the other distributions, \( LE(x) \) would be equal to it). Proximity to such a distribution would be of interest in evaluating each of the available distributions at the first order imperative. Hence we are led to contemplate,

\[
\min_{M(x)} LEPB(M(x)) = -\frac{1}{\mu} \int_{0}^{\infty} [M(x) - LE(x)] dx
\]  

(5)

where \( M(x) = \{G(x), H(x), J(x), \ldots, K(x)\} \), since the lowest value of \( LEPB \) represents the closest proximity to the envelope.

If the policymaker’s imperative is represented by the \( j \)’th degree dominance criterion, one could contemplate the lower frontier or envelope of all possible \( j \)’th order integrals of the candidate distributions, \( G_j(x), H_j(x), J_j(x), \ldots, K_j(x) \), which would be \( LE_j(x) = \min\{G_j(x), H_j(x), J_j(x), \ldots, K_j(x)\} \). Note that \( LE_j(x) \) would stochastically dominate all distributions \( G(x), H(x), J(x), \ldots, K(x) \), at the \( j \)’th order and would thus, if it existed, be the preferred distribution at that order\(^8\). Proximity to such a distribution would be of interest in evaluating the available distributions, hence she would be led to contemplate,

\[
\min_{M_j(x)} LEPB_j(M_j(x)) = -\frac{1}{\mu_j} \int_{0}^{\infty} [M_j(x) - LE_j(x)] dx
\]  

(6)

where \( M_j(x) = \{G_j(x), H_j(x), J_j(x), \ldots, K_j(x)\} \)

\(^8\)Again note that if one of the distributions \( j \)th Order Dominated all of the other distributions \( LE_j(x) \) would be equal to it.
4 An Illustration

To illustrate these ideas, we contemplate three alternative non-combinable policies, A, B, and C that yield the same per capita return in terms of expected post-tax income to society. The different policies have different redistributive effects, which will be characterized through different revenue neutral tax policies on the initial distribution, which shall be denoted policy A. Using the results of Moyes and Shorrocks (1994) it is assumed that the effect of policy B is that of a proportionate tax $t_p(x) = t$, where $0 < t < 1$, whose aggregate proceeds are distributed equally across the population at a level of $M$ per person. The effect of policy C was equivalent to a progressive tax $t_{pr}(x) = t_1 + t_2F(x)$ (where $F(x)$ is the cumulative distribution of $f(x)$, the income size distribution of pre-tax income $x$, and $0 < t_1 + t_2F(x) < 1$, so that $0 < t_1 < 1$, and $0 < t_2 < 1 - t_1$), and again the aggregate per capita proceeds $M$ is distributed equally across the population. All tax regimes are revenue neutral, which implies that the post-tax income for policy B is $(1 - t)x + M$, and revenue neutrality implies:

$$\int (tx - M)dF(x) = 0 \Rightarrow M = tE(x)$$

(7)

and for policy C, post-tax income will be $(1 - t_1 - t_2F(x))x + M$, with revenue neutrality implying:

$$\int ((t_1 + t_2F(x))x - M)dF(x) = 0 \Rightarrow t_2 = \left(\frac{M - t_1E(x)}{\int xF(x)dF(x)}\right)$$

(8)

The empirical analogues of the policies applied to a random sample of $n$ pre-tax weekly incomes $x_i$, $i = 1, \ldots, n$ (where incomes $x$ are ranked highest 1 to lowest $n$) drawn from the Canadian Labour Force Survey for January 2012 (wage rate multiplied by usual hours of work per week) would yield post-policy incomes $y_i$ of,

- **A**: $y_i = x_i$
- **B**: $y_i = (1 - t)x_i + M$
- **C**: $y_i = \left[1 - t_1 - t_2\left(1 - \frac{\text{rank}(x_i)}{n}\right)\right] x_i + M$

Income distributions that are the result of the three policy alternatives are illustrated in figure 1. The sample size was 52,173, the parameters were chosen as $t = 0.5$, $t_1 = 0.3$, and as a consequence $t_2 = 0.2976$, and summary statistics for the three policies are presented.
Figure 1: Density Functions of Policy Outcomes

Table 1: Income Distribution Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Policy A</th>
<th>Policy B</th>
<th>Policy C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Income</td>
<td>836.89</td>
<td>836.89</td>
<td>836.89</td>
</tr>
<tr>
<td>Median Income</td>
<td>750.00</td>
<td>793.44</td>
<td>831.73</td>
</tr>
<tr>
<td>Standard Deviations</td>
<td>534.41</td>
<td>267.20</td>
<td>205.10</td>
</tr>
<tr>
<td>Maximum Income</td>
<td>5769.60</td>
<td>3303.24</td>
<td>2740.27</td>
</tr>
<tr>
<td>Minimum Income</td>
<td>4.80</td>
<td>420.84</td>
<td>421.80</td>
</tr>
</tbody>
</table>

in Table 1. All three distributions have the same average income with the dispersion ranking $A > B > C$, all are right skewed with Policy $C$ being the least skewed.

Table 2 reports the dominance relationships between the policies in terms of the maximum and minimum differences between the 1st, 2nd, and 3rd orders of integration of the respective distributions (positive maximums together with negative minimums imply no dominance relationship at that order of integration). As is evident, there are no dominance relationships between the policy outcomes at the 1st order, at the 2nd order $A$ is dominated by both $B$ and $C$, though there is no dominance relationship between $B$ and $C$, and at the 3rd order comparison outcome $C$ universally dominates, and will be the
envelope of the three distributions at that level of dominance comparison. Note incidentally that if a Rawlsian, infinite order dominance, imperative confronted the policymaker, policy C would be the choice since it presents the best outcome for the poorest person. Nonetheless, the primary point here is stochastic dominance’s inability to provide a resolution to the policy choice problem should the policymaker’s imperative be utilitarian, or Daltonian in nature.

Table 2: Between Policy Dominance Comparisons \((A \succ_k B)\) implies \(k^{th}\) order dominance of \(A\) over \(B\)

<table>
<thead>
<tr>
<th></th>
<th>(A - B)</th>
<th>(A - C)</th>
<th>(B - C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(^{st}) Order minimum difference</td>
<td>-0.1272</td>
<td>-0.1844</td>
<td>-0.0734</td>
</tr>
<tr>
<td>1(^{st}) Order maximum difference</td>
<td>0.2359</td>
<td>0.2527</td>
<td>0.1019</td>
</tr>
<tr>
<td>2(^{nd}) Order minimum difference</td>
<td>0.0000</td>
<td>0.0000</td>
<td>(-1.9895e^{-12})</td>
</tr>
<tr>
<td>2(^{nd}) Order maximum difference</td>
<td>0.1225</td>
<td>0.1531</td>
<td>0.0318</td>
</tr>
<tr>
<td>3(^{rd}) Order minimum difference</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>3(^{rd}) Order maximum difference</td>
<td>0.1529</td>
<td>0.1738</td>
<td>0.0209</td>
</tr>
</tbody>
</table>

To highlight the merit of this comparison technique, consider the Non-Standardized Policy Evaluation Indices reported in Table 3. Under a Utilitarian imperative, Policy A would be chosen (although the magnitudes of each respective policy index suggests that there is very little to choose between the policies at this order of dominance comparison). Under a second order inequality averse imperative, Policy C would be chosen, and under a third order inequality averse imperative, where poorer agents are of greater concern, Policy C would still be chosen (note here the Index is zero because Policy C’s distribution, in being uniformly dominant at the third order over the other distributions, will constitute the lower envelope at that order).
Table 3: Policy Evaluation Indices

<table>
<thead>
<tr>
<th></th>
<th>Policy A</th>
<th>Policy B</th>
<th>Policy C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{LEPB}_1(\mathcal{M}(x))$</td>
<td>22.2199</td>
<td>22.2359</td>
<td>22.2310</td>
</tr>
<tr>
<td>$\mathcal{LEPB}_2(\mathcal{M}(x))$</td>
<td>25.2379</td>
<td>3.0398</td>
<td>1.4205e^{-10}</td>
</tr>
<tr>
<td>$\mathcal{LEPB}_3(\mathcal{M}(x))$</td>
<td>140.9837</td>
<td>17.1928</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

References


